B41

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12Q 1/68

A2

(11) International Publication Number: WO 00/61808

(43) International Publication Date: 19 October 2000 (19.10.00)

(21) International Application Number: PCT/US00/09263

(22) International Filing Date: 7 April 2000 (07.04.00)

(30) Priority Data:
60/128,629 9 April 1999 (09.04.99) US

(71) Applicant: EXACT LABORATORIES, INC. [US/US]; 63 Great Road, Maynard, MA 01754 (US).

(72) Inventors: LAPIDUS, Stanley, N.; 12 Old Evergreen Road, Bedford, NH 03110 (US). SHUBER, Anthony, P.; 11 Grant Street, Milford, MA 01757 (US).

(74) Agent: MEYERS, Thomas, C.; Testa, Hurwitz & Thibeault, LLP, High Street Tower, 125 High Street, Boston, MA 02110 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: METHODS FOR DETECTING NUCLEIC ACIDS INDICATIVE OF CANCER

(57) Abstract

The invention provides methods for screening tissue or body fluid samples for nucleic acid indicia of cancer or precancer.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

METHODS FOR DETECTING NUCLEIC ACIDS INDICATIVE OF CANCER

FIELD OF THE INVENTION

This invention relates to methods for the early detection of cancer in patients by screening for large DNA fragments. Methods of the invention are especially useful in the detection of colon cancer.

BACKGROUND OF THE INVENTION

Alterations in genomic integrity often are associated with disease or with the propensity for disease. For example, many cancers are thought to arise through a series of mutations in genomic DNA, resulting in genomic instability in the form of uncontrolled cellular growth. In normal cells, damage to genomic DNA typically leads to expression of tumor suppressors, such as the cell-cycle regulator, p53. For example, damage to cellular DNA results in increased expression of p53 which arrests the cell cycle to allow repair of the damage. If the damaged DNA cannot be repaired, the cell undergoes apoptosis, thus preventing the accumulation of additional mutations in daughter cells. If however, there is a mutation in the p53 gene itself (or in another cell cycle regulator), damaged cells will proceed through the cell cycle, giving rise to progeny in which additional DNA mutations will go unchecked. It is the accumulation of these mutations that is the hallmark of many cancers.

The process of apoptosis is important not only in the regulation of cellular metabolism, but also in inhibiting oncogenesis. As cells undergo apoptosis, the nucleus becomes small and fragmented. Nuclear DNA is digested into spindle fragments that are generally no larger than about 200 base pairs. As the process continues, usually through multiple pathways, the cell membrane breaks down, and cellular contents are metabolized. As a result, cells that have the potential to enter the multi-step pathway leading to cancer are eliminated.

Many cancers are curable if detected early in their development. For example, colorectal cancers typically originate in the colonic epithelium, and are not extensively vascularized (and therefore not invasive) during early stages of development. The transition to a highly-vascularized, invasive and ultimately metastatic cancer commonly

takes ten years or longer. If the presence of cancer is detected prior to extensive vascularization, surgical removal typically is an effective cure. However, colorectal cancer is often detected only upon manifestation of clinical symptoms, such as pain and black tarry stool. Generally, such symptoms are present only when the disease is well established, and often after metastasis has occurred. Early detection of colorectal cancer therefore is important in order to significantly reduce its morbidity.

Invasive diagnostic methods, such as endoscopic examination, allow direct visual identification, removal, and biopsy of potentially-cancerous tissue. Endoscopy is expensive, uncomfortable, inherently risky, and not a practical tool for early diagnosis.

Established non-invasive screening methods involve assaying stool samples for the presence of fecal occult blood or for elevated levels of carcinoembryonic antigen, both of which are suggestive of the presence of colorectal cancer. Additionally, recent developments in molecular biology provide methods of great potential for detecting the presence of a range of DNA mutations indicative of colorectal cancer. The presence of such mutations can be detected in DNA found in stool samples during various stages of colorectal cancer. However, stool comprises cells and cellular debris from the patient, from microorganisms, and from food, resulting in a heterogeneous population of cells. This makes detection of small, specific subpopulations difficult to detect reliably.

There is a need in the art for additional non-invasive methods for early diagnosis of cancer that will detect characteristics indicative of the presence of cancer.

SUMMARY OF THE INVENTION

The present invention provides methods for identifying indicia of cancer in tissue or body fluid samples by identifying non-apoptotic DNA in those samples. The invention also provides methods for identifying indicia of cancer or precancer in samples containing exfoliated epithelial cells. It has now been recognized that DNA obtained from exfoliated normal (non-cancerous) cells is different than DNA obtained from exfoliated cancer or precancer cells. Normal exfoliated cells typically have undergone apoptosis, and thus produce cells or cellular debris (depending upon the stage of apoptosis) comprising DNA that has been substantially degraded. Exfoliated cancer or precancer cells typically have not undergone apoptosis, and such cells or their debris, while producing some very small fragments as a result of degradation in the sample, typically also contain a higher proportion of large DNA fragments

(compared to those observed in cells or debris from exfoliated normal cells). The difference in DNA integrity between normal and abnormal cells is a marker for the presence of cancer or precancer in a sample comprising exfoliated cells.

Stool is a good sample for exemplification of methods of the invention. The colonic epithelium undergoes a continual process of exfoliation. Normal epithelial cells undergo apoptosis, and are sloughed into the lumen of the colon, and onto forming stool. Cells from polyps and tumors are also sloughed onto forming stool. However, cells from polyps or tumors are, by definition, not apoptotic. Methods of the invention take advantage of the different characteristics between apoptotic and non- apoptotic cells in order to screen patient samples for indicia of cancer or precancer.

As noted above, non-cancerous (normal) cells undergo apoptosis at regular intervals, or in response to irreparable cell damage. As a result of apoptosis, DNA from normal cells is cleaved into small fragments having about 200 or fewer base pairs, and typically 180 base pairs or less. In contrast, DNA obtained from cancer or precancer cells is much larger than the typical apoptotic fragments. Thus, the presence of large DNA fragments in a sample (e.g., of sloughed colonic epithelium) indicates that there are or were cells in the sample (or the specimen from which it was obtained) that have avoided apoptosis, and its coincidental degradation of DNA. The presence of large DNA fragments represents a positive screen for cancer or precancer.

Accordingly, methods of the invention comprise detecting the presence in a biological sample of species-specific nucleic acids indicative of cancer or precaner. Samples comprising such nucleic acids are identified as having indicia of cancer or precancer. In preferred methods, patients presenting samples having a high proportion of non-apoptotic nucleic acids as determined by methods of the invention are further evaluated for the presence of a tumor, adenoma, or other cancerous or precancerous lesion.

In general, methods of the invention comprise detecting in a biological sample one or more DNA fragment(s) of a length that would not be substantially present in noncancerous cells or cellular debris. In a preferred embodiment, such fragments are larger than a typical apoptotic spindle fragment, or larger than about 170 base pairs. However, also in a preferred embodiment, methods of the invention comprise detecting DNA fragments that are greater than about 200 base pairs, and preferably greater than about 500 base pairs. There is no upper limit on these fragments, as all that is

necessary is that the fragment be larger than an apoptotic fragment. Typically, however, fragments indicative of cancer or precancer cells are between about 200 and about 3500 base pairs, and ideally between about 500 and about 2500 base pairs.

Accordingly, in a preferred embodiment, methods of the invention comprise detecting in a tissue or body fluid sample the presence of nucleic acid fragments having greater than about 500 base pairs or having a molecular weight corresponding to greater than about 500 base pairs. In other preferred embodiments, methods of the invention comprise detecting nucleic acid fragments having between about 200 and about 1000 base pairs, preferably between about 200 and about 600 base pairs, and most preferably about 500 base pairs.

Also in a preferred embodiment, methods of the invention comprise determining a ratio of large fragments (200-3500 bp) to small fragments (less than 200 bp), and determining whether the ratio exceeds an empirically-derived threshold. The threshold is determined empirically by analyzing ratios of large-to-small fragments, and correlating those ratios with the disease state of a selected population of normal and cancer patients. In preferred embodiments, amounts of large and small fragments are determined by polymerase chain reaction amplification of sample DNA using primers selected to amplify long and short fragments. Alternatively, amounts of large and small fragments are determined using the same primer.

Preferred methods of the invention comprise amplifying nucleic acids in a representative stool sample using human-specific primers, and detecting amplicons having greater than about 200, and preferably about 500 or more base pairs. In a highly-preferred embodiment, amplification is accomplished by polymerase chain reaction (PCR) using forward and reverse primers directed against human-specific nucleic acid fragments, and spaced apart to provide a lower limit on the resulting amplicons. Also in a highly-preferred embodiment, primers for PCR are directed against human oncogene or tumor suppressor sequences. Preferred target nucleic acids for PCR primers include p53, Kras, apc, dcc, and other genes known or suspected to be associated with cancer, and especially colorectal cancer. Methods for conducting PCR are provided in U.S. Patent No. 4,683,202, incorporated by reference herein. The presence of amplicon greater than about 200 base pairs in length is indicative of template nucleic acid in the sample of that length (or longer). According to methods of the invention such long sequences represent a positive screen, and are

indicative of cancer or precancer.

Preferred biological samples include stool, pus and urine. Method of the inventions are especially useful for the detection of large DNA fragments in samples comprising exfoliate. Tissue (e.g., colon, lungs, bladder) in which cells, especially epithelial cells, are exfoliated are most preferred for screening methods of the invention. In such tissues, continuing cellular renewal requires that cells are regularly sloughed after having undergone apoptosis. Samples of the exfoliate (tissue or body fluid containing the exfoliated cells) predominantly comprise apoptotic DNA.

Preferred methods of the invention for use on a stool sample comprise obtaining a representative stool sample. An especially-preferred method for preparing a stool sample is disclosed in U.S. Patent No. 5,741,650, and in co-owned, co-pending U.S. patent application No. 09/059,713 (attorney docket No. EXT-015), each of which is incorporated by reference herein.

In a preferred embodiment, methods of the invention comprise homogenizing a representative stool sample in a solvent in order to form a homogenized sample mixture having a solvent volume to stool mass ratio of at least 5 to 1. An especially-preferred ratio of solvent volume to stool mass is about 20:1. A preferred solvent for preparing stool samples according to the invention is a physiologically-compatible buffer comprising a detergent and a proteinase and optionally a DNase inhibitor, such as a buffer comprising Tris-EDTA-NaCl. A preferred buffer is 50 mM Tris, 150 mM EDTA and 10 mM NaCl at pH 9.0. Another preferred solvent is guanidine isothiocyanate (GITC). Providing an optimal solvent volume to stool mass ratio increases the yield of nucleic acid generally from the sample. Further details regarding sample preparation are disclosed in co-owned, co-pending U.S. Patent Application No. 08/XXXXXXX (Attorney Docket No. EXT-0028) incorporated herein by reference.

Preferred methods of the invention further comprise enriching sample for human DNA. Preferred enrichment methods for use in the invention include enriching a desired human target sequence using an affinity column, sequence-specific capture, or through the use of preferred buffers that bias isolation of human DNA. A preferred enrichment method is based upon the capture of unique human nucleic acids using, for example, an affinity column. Details of such methods are provided below.

In a preferred embodiment, methods further comprise the step of extracting DNA from the homogenized sample mixture using sequence-specific nucleic acid probes.

Particularly preferred are probes hybridizing to human DNA. The probes are preferably labeled. Preferred labels include radioactive labels, fluorescent labels, molecular weight labels and enzymatic labels. Other labels are well known in the art.

In a preferred embodiment gel electrophoresis, affinity chromatography, or mass spectrometry are used to detect large DNA fragments (fragments comprising greater than about 200 base pairs). The presence of large DNA fragments in the sample is indicative of colorectal cancer.

In a preferred embodiment capture probes comprise DNA, RNA or PNA, and are detectably labeled using methods known in the art. In one embodiment probes are labeled with radioactive isotopes such as ^{32}P, ^{33}P, ^{35}S, ^{125}I, or any other detectable isotope useful for labeling a hybridization probe. In an another embodiment, probes are labeled with fluorescent molecules. Numerous fluorescent labels are known in the art, and any detectable fluorescent probe label is useful for practice of the invention. Alternatively, probes are attached to moieties which increase their molecular weight. For example a probe may be directly attached to a glycoprotein, or a glass bead, or any compound which has a detectable effect on the molecular weight of the probe. In a further embodiment, probes are labeled with a compound that is detectable due to specific interactions with an additional compound. For example, biotinylated probes are detectable via interaction with streptavidin. The streptavidin moiety is attached to a detectable label such as a bead, a fluorescent tag, or an enzyme. In another example, the probes are labeled with a hapten or other antigen which is specifically recognized by an antibody. The antibody is made detectable using methods known in the art including radioactive isotopes, fluorescent tags, and enzyme reactions. In a further example the probes are directly attached to an enzyme which is detectable via a specific enzyme catalyzed reaction generating a detectable product.

Finally, methods of the invention allow one to approximate the position in the colon of a colorectal lesion based upon the relative amount of DNA fragments in a stool sample that are greater than 200 base pairs in length. This aspect of the invention relies on the fact that the lytic properties of stool are greater in the proximal colon than they are in the distal colon. In the proximal colon, stool is typically in liquid form. Therefore, the cell lysis and DNA degrading enzymes in the colon have greater access to exfoliated cells in the liquid mixture of the proximal colon as compared to their access to exfoliated cells sloughed onto formed or forming stool that is typical in the

distal colon. As a consequence of the differences between the environments of the proximal and distal colon, the present invention provides that typical DNA fragments from cells exfoliated into the proximal colon are smaller than DNA fragments from cells exfoliated into the distal colon. Figure 1 provides an example of the progression of DNA sizes expected for cancer or precancer cells exfoliated into different regions of the colon. The size of DNA fragments from non-cancerous or precancerous cells is the same throughout the colon due to the fact that the DNA from those cells is degraded primarily through apoptosis. Thus, cell lysis and DNA degradation play only minor roles in determining the size of DNA fragments from most exfoliated normal cells throughout the colon. It is noted, however, that normal cells that are, for example, mechanically sheared from the colon undergo the same lytic and degradation cycle as the typical cancer or precancer cell. However, the contribution of such non-apoptotic normal cells to the overall level of DNA in the stool sample is small, and is controlled for by establishing standards as taught below.

Further aspects and advantages of the invention are contained in the following detailed description thereof.

Description of the Drawings

Figure 1 shows a schematic representation of the colon, and the representative (typical) DNA fragment length for DNA obtained from a cancer or precancer exfoliated cell over the representative (typical) apoptotic (normal) DNA fragment length for various regions of the colon.

Figure 2 is a gel photograph showing results of amplification of Kras (exon 1) DNA isolated from stool using forward and reverse primers spaced about 200 bp apart. The band intensity relates to the amount of 200bp product or greater in the sample. Lanes 1-4 are results from patients with cancer or adenoma, lane 5 is a positive control, lanes 6-10 are from patients who did not have cancer or adenoma, lanes 11-12 are negative controls, and lanes 13-18 are standards at the approximate molecular weight indicated in the figure.

Figure 3 is a gel photograph showing results of amplification of apc (exon 15) DNA isolated from stool using forward and reverse primers spaced about 200 bp apart. The band intensity relates to the amount of 200bp product or greater in the sample. Lanes 1-4 are results from patients with cancer or adenoma, lane 5 is a positive control,

lanes 6-10 are from patients who did not have cancer or adenoma, lanes 11-12 are negative controls, and lanes 13-18 are standards at the approximate molecular weight indicated in the figure.

Figure 4 is a gel photograph showing results of amplification of apc (exon 15) DNA isolated from stool using forward and reverse primers spaced about 200 bp apart. The band intensity relates to the amount of 200bp product or greater in the sample. Lanes 1-4 are results from patients with cancer or adenoma, lane 5 is a positive control, lanes 6-10 are from patients who did not have cancer or adenoma, lanes 11-12 are negative controls, and lanes 13-18 are standards at the approximate molecular weight indicated in the figure.

Figure 5 is a gel photograph showing results of amplification of apc (exon 15) DNA isolated from stool using forward and reverse primers spaced about 200 bp apart. The band intensity relates to the amount of 200bp product or greater in the sample. Lanes 1-4 are results from patients with cancer or adenoma, lane 5 is a positive control, lanes 6-10 are from patients who did not have cancer or adenoma, lanes 11-12 are negative controls, and lanes 13-18 are standards at the approximate molecular weight indicated in the figure.

Figure 6 is a gel photograph showing results of amplification of p53 (exon 5) DNA isolated from stool using forward and reverse primers spaced about 200 bp apart. The band intensity relates to the amount of 200bp product or greater in the sample. Lanes 1-4 are results from patients with cancer or adenoma, lane 5 is a positive control, lanes 6-10 are from patients who did not have cancer or adenoma, lanes 11-12 are negative controls, and lanes 13-18 are standards at the approximate molecular weight indicated in the figure.

Figure 7 is a gel photograph showing results of amplification of p53 (exon 7) DNA isolated from stool using forward and reverse primers spaced about 200 bp apart. The band intensity relates to the amount of 200bp product or greater in the sample. Lanes 1-4 are results from patients with cancer or adenoma, lane 5 is a positive control, lanes 6-10 are from patients who did not have cancer or adenoma, lanes 11-12 are negative controls, and lanes 13-18 are standards at the approximate molecular weight indicated in the figure.

Figure 8 is a gel photograph showing results of amplification of p53 (exon 8) DNA isolated from stool using forward and reverse primers spaced about 200 bp apart.

The band intensity relates to the amount of 200bp product or greater in the sample. Lanes 1-4 are results from patients with cancer or adenoma, lane 5 is a positive control, lanes 6-10 are from patients who did not have cancer or adenoma, lanes 11-12 are negative controls, and lanes 13-18 are standards at the approximate molecular weight indicated in the figure.

Figure 9 is a gel photograph of results of amplification of DNA from stool samples using forward and reverse primers spaced approximately 1.8Kb apart. The band intensity shows the amount of 1.8Kb or greater product. Lanes 1, 8, and 9 are negative controls, lanes 2, 3, and 5 are results from patients with cancer or adenoma, lanes 4, 6, and 7 are results from patients who did not have cancer or adenoma, and lanes 10-14 are molecular weight standards.

Figure 10 is a gel photograph of results of amplification of DNA from stool samples using forward and reverse primers spaced approximately 1.8Kb apart. The band intensity shows the amount of 1.8Kb or greater product. Lanes 1, 8, and 9 are negative controls, lanes 2, 3, and 5 are results from patients with cancer or adenoma, lanes 4, 6, and 7 are results from patients who did not have cancer or adenoma, and lanes 10-14 are molecular weight standards.

Figure 11 is a gel photograph of results of amplification of DNA from stool samples using forward and reverse primers spaced approximately 1.8Kb apart. The band intensity shows the amount of 1.8Kb or greater product. Lanes 1, 8, and 9 are negative controls, lanes 2, 3, and 5 are results from patients with cancer or adenoma, lanes 4, 6, and 7 are results from patients who did not have cancer or adenoma, and lanes 10-14 are molecular weight standards.

DETAILED DESCRIPTION OF THE INVENTION

Methods of the invention are based upon the observation that samples comprising cells from patients with cancer or precancer contain a greater amount of high molecular weight (long sequence) DNA fragments as compared to corresponding samples obtained from individuals that are free of cancer/precancer. Accordingly, methods of the invention provide accurate screening and diagnostic procedures for cancer or precancer.

Methods of the invention are useful to detect nucleic acid indicia of cancer or precancer in any tissue or body fluid sample. For example, sputum samples are used

to detect the presence of high molecular weight (long sequence) DNA as a marker for cancer. The majority of cells exfoliated into sputum have undergone apoptosis and subsequent further enzymatic degradation. The predominant DNA from those cells is small, apoptotic DNA. Cancer cells produced by, for example, the lungs, the nasal passages, or the trachea will also be sloughed into sputum. However, the DNA from those cells, while being exposed to enzymatic processes, has not been affected by apoptosis. Accordingly, fragments from cancer or precancer cells found in sputum are larger than fragments expected to be produced by normal cells.

Similarly, cells sloughed by cancerous or precancerous lesions in the bladder or kidney produce non-apoptotic DNA in urine, cancerous or precancerous lesions in the lymph nodes result in non-apoptotic DNA fragments in lymph, and cancerous or precancerous cells in the breast slough non-apoptotic DNA-containing cells that can be harvested via aspiration. Accordingly, methods of the invention are useful in any tissue or body fluid. However, for purposes of exemplification of the methods described herein, stool sample were used to predict the presence of colorectal cancer or precancer. Stool is an excellent specimen for analysis due to the characteristic exfoliation of colonic epithelia as described above.

Methods of the invention are practiced by detecting the presence of DNA fragments having a sequence length that would not be expected to be present in significant amounts in a sample obtained from a healthy individual (i.e., an individual who does not have cancer or precancer). A threshold amount of large fragments is an amount that exceeds a predetermined level expected or determined for non-cancerous/non-precancerous cells. The predetermined level or standard can be determined by detecting the amount of a particular size of DNA fragment (preferably apoptotic fragments characteristic of normal cells) in a population or subpopulation of normal patients. Standards can be determined empirically, and, once determined, can be used as the basis for further screening.

The size of fragments to be used is chosen based upon the convenience of the individual performing the screen. Factors affecting the size of fragments used in screening or diagnostic methods of the invention include the availability and costs of probes and primers, the desired target of amplification, the type of cancer being screened, and the patient sample on which screening takes place. The invention takes advantage of the recognition that large fragments exist in greater abundance in

abnormal samples than in normal samples. Accordingly, the precise size of fragments used in methods of the invention does not matter. For any given size of fragments to be analyzed, a cutoff must be determined to distinguish between normal and abnormal samples. Preferably, the cutoff is determined empirically based upon known normal and abnormal sample, and then is used in future screenings.

The following examples provide further details of methods according to the invention. For purposes of exemplification, the following examples provide details of the use of the method if the present invention in colon cancer detection. Accordingly, while exemplified in the following manner, the invention is not so limited and the skilled artisan will appreciate its wide range of application upon consideration thereof.

Exemplary Method for the Detection of Colon Cancer

For the analysis of stool samples, preferred methods of the invention comprise obtaining at least a cross-section or circumfrential portion of a voided stool as taught in U.S. patent number 5,741,650, and co-pending, co-owned U.S. patent application serial number 09/059,718, both of which are incorporated by reference herein. While a cross-sectional or circumfrential portion of stool is desirable, methods provided herein are conducted on random samples obtained from voided stool, which include smears or scrapings. Once obtained, the stool specimen is homogenized. A preferable buffer for homogenization is one that contains at least 16mM ethylenediaminetetraacetic acid (EDTA). However, as taught in co-pending, co-owned U.S. patent application serial number 60/122,177, incorporated by reference herein, it has been discovered that the use of at least 150mM EDTA greatly improves the yield of nucleic acid from stool. Thus, a preferred buffer for stool homogenization comprises phosphate buffered saline, 20-100 mM NaCl or KCl, at least 150mM EDTA, and optionally a detergent (such as SDS) and a proteinase (e.g., proteinase K).

After homogenization, nucleic acid is preferably isolated from the stool sample. Isolation or extraction of nucleic acid is not required in all methods of the invention, as certain detection techniques can be adequately performed in homogenized stool without isolation of nucleic acids. In a preferred embodiment, however, homogenized stool is spun to create a supernatant containing nucleic acids, proteins, lipids, and other cellular debris. The supernatant is treated with a detergent and proteinase to degrade protein, and the nucleic acid is phenol-chloroform extracted. The extracted nucleic

acids are then precipitated with alcohol. Other techniques can be used to isolate nucleic acid from the sample. Such techniques include hybrid capture, and amplification directly from the homogenized stool. Nucleic acids can be purified and/or isolated to the extent required by the screening assay to be employed. Total DNA is isolated using techniques known in the art.

Once DNA is isolated, the sample preferably is enriched for human nucleic acids using sequence specific capture probes. Pelletized DNA is resuspended in TE buffer. Guanidine isothiocyanatate (GITC) is then added. An excess of capture probes that target human DNA are added to the sample. The sample is heated to denature the DNA and then cooled. Finally, probe and target DNA are allowed to hybridize. Steptavidin -coated magnetized beads are suspended in water and added to the mixture. After briefly mixing, the mixture is maintained at room temperature for approximately 30 minutes. Once the affinity binding is completed, a magnetic filed is applied to the sample to draw the magnetized isolation beads (both with and without hybridized complex). The beads are then washed four (4) times in 1M GITC/0.1% Igepal (Sigma, St. Louis, MO) solution for 15 minutes, followed by two (2) washes with warm buffer (TE with 1M NaCl) for 15 minutes in order to isolate complexed streptavidin. Finally, distilled water is added to the beads and heated to elude the DNA. Gel electrophoresis can then be performed on the human DNA that has been captured.

III. Determination of Fragment Length

The size of human DNA fragments obtained above can be determined by numerous means. For example, human DNA can be separated using gel electrophoresis. A 5% acrylamide gel is prepared using techniques known in the art. See Ausubel et. al., Short Protocols in Molecular Biology, John Wiley & Sones, 1195, pgs. 2-23-2-24, incorporated by reference herein. The size of human DNA fragments is then determined by comparison to known standards. Fragments greater than about 200 bp provide a positive screen. While a diagnosis can be made on the basis of the screen alone, patients presenting a positive screen are preferably advised to seek follow-up testing to render a confirmed diagnosis.

A preferred means for determining human DNA fragment length is by using PCR. Methods for implementing PCR are well-known. In the present invention, human DNA fragments are amplified using human-specific primers. Amplicon of greater than

about 200 bp produced by PCR represents a positive screen. Other amplification reactions and modifications of PCR, such as ligase chain reaction, reverse-phase PCR, Q-PCR, and others may be used to produce detectable levels of amplicon. Amplicon may be detected by coupling to a reporter (e.g. flouresence, radioisotopes, and the like), by sequencing, by gel electrophoresis, by mass spectrometry, or by any other means known in the art, as long as the length, weight, or other characteristic of the amplicons identifies them by size.

EXAMPLES

Experiments were conducted to determine whether characteristics of amplifiable DNA in stool were predictive of cancer or precancer in patients from whom stools samples were obtained. In the first experiment, the amount of amplifiable DNA was measured in each of several stool samples using PCR amplification to detect DNA fragments in the sample of at least 200 base pairs in length. The second experiment determined the amount of long (greater than 200 base pair) fragments in the same samples, and then to determine ratios of long product to short product.

I. The Use of Amplifiable DNA as a Marker for Cancer or Precancer

Stool samples were collected from 9 patients who presented with symptoms or a medical history that indicated that a colonoscopy should be performed. Each stool sample was frozen. Immediately after providing a stool sample, each patient was given a colonoscopy in order to determine the patient's disease status. Based upon the colonoscopy results, and subsequent histological analysis of biopsy samples taken during colonsocopy, individuals were placed into one of two groups: normal or abnormal. The abnormal group consisted of patients with cancer or with an adenoma of at least 1 cm in diameter. Based upon these results, 4 of the 9 patients were placed into the abnormal group.

The samples were screened by hybrid capturing human DNA, and determining the amount of amplifiable DNA having at least 200 base pairs. Each frozen stool specimen, weighing from 7-33 grams, was though and homogenized in 500 mM Tris, 16 mM EDTA, and 10 mM NaCl, pH 9.0 at a volume: to mass ratio of 3:1. Samples were then rehomogenized in the same buffer to a final volume-to-mass ratio of 20:1, and spun in glass macro beads at 2356 x g. The supernatant was collected and treated with SDS and proteinase k. The DNA was then phenol-chloroform

extracted and precipitated with alcohol. The precipitate was suspended in 10 mM Tris and 1 mM EDTA (1 x TE), pH 7.4. Finally, the DNA was treated with Rnase.

Human DNA was isolated from the precipitate by sequence-specific hybrid capture. Biotynilated probes against portions of the p53, K-ras, and apc genes were used. The K-ras probe was 5'GTGGAGTATTTGATAGTGTATTAACCTTATGTGTGAC 3' (SEQ ID NO: 1). There were two apc probes: apc-1309 was 5'TTCCAGCAGTGTCACAGCACCCTAGAACCAAATCCAG 3' (SEQ ID NO: 2), and apc-1378 was 5'CAGATAGCCCTGGACAAACAATGCCACGAAGCAGAAG 3' (SEQ ID NO: 3). There were four probes against p53, the first (hybridizing to a portion of exon 5) was 5'TACTCCCCTGCCCTCAACAAGATGTTTTGCCAACTGG3' (SEQ ID NO:4), the second (hybridizing to a portion of exon 7) was 5'ATTTCTTCCATACTACTACCCATCGACCTCTCATC3' (SEQ ID NO: 5), the third, also hybridizing to a portion of exon 7 was 5'ATGAGGCCAGTGCGCCTTGGGGAGACCTGTGGCAAGC3' (SEQ ID NO: 6); and finally, a probe against exon 8 had the sequence 5'GAAAGGACAAGGGTGGTTGGGAGTAGATGGAGCCTGG3' (SEQ ID NO: 7). A 10 ul aliquot of each probe (20 pmol/capture) was added to a suspension containing 300 ul DNA in the presence of 310ul 6M GITC buffer for 2 hours at room temperature. Hybrid complexes were isolated using streptavidin-coated beads (Dynal). After washing, probe-bead complexes were suspended at 25° C for 1 hour in 0.1x TE buffer, pH7.4. The suspension was then heated for 4 minutes at 85° C, and the beads were removed.

Captured DNA was then amplified using PCR, essentially as described in U.S. Patent No. 4,683,202, incorporated by reference herein. Each sample was amplified using forward and reverse primers through 7 loci (Kras, exon 1, APC exon 15 (3 separate loci), p53, exon 5, p53, exon 7, and p53, exon 8) in duplicate (for a total of 14 amplifications for each locus). Seven separate PCRs (33 cycles each) were run in duplicate using primers directed to detect fragments in the sample having 200 base pairs or more. Amplified DNA was placed on a 4% Nusieve (FMC Biochemical) gel (3% Nusieve, 1% agarose), and stained with ethidium bromide (0.5 ug/ml). The resulting amplified DNA was graded based upon the relative intensity of the stained gels. The results are shown in Figures 2-8. Each Figure represents the results for all 9 patients (including standards) for the seven different loci that were amplified. As shown in the Figures, each sample from a patient with cancer or adenoma was detected as a band

having significantly greater intensity than the bands associated with samples from patients who did not have cancer or precancer. All four cancer/adenoma patients identified using colonoscopy were correctly identified by determining the amount of amplifiable DNA 200 base pairs or greater in length. As shown in Figures 2-8, the results were the same regardless of which locus was amplified. Accordingly, the amount of 200bp or greater DNA in a sample was predictive of patient disease status.

EXAMPLE 2

An experiment was conducted that was essentially identical to the one described above in Example 1, but forward and reverse primers were placed such that fragments of about 1.8 Kb and above were amplified.

DNA was prepared as described above. Forward and reverse primers were spaced so as to hybridize approximately 1.8Kb apart on three different loci (Kras, exon 1, APC, exon 15, and p53 exon 5). Thirty-three rounds of amplification were performed, and the resulting DNA was placed on a 5% acrylamide gel. The results are shown in Figures 9-11. As shown in the Figures (which show results from three separate experiments to amplify and detect "long" product), samples from individuals having cancer or precancer produced large amounts of high-molecular weight (in this case 1.8Kb and above) DNA; whereas samples from patients who did not have cancer or precancer produced no DNA in the range of about 1.8Kb and higher. Thus, the presence of high-molecular weight DNA was indicative of the disease status of the patient.

The invention has been described in terms of its preferred embodiments. Alternative embodiments are apparent to the skilled artisan upon examination of the specification and claims.

CLAIMS

What is claimed is:

1. A method for screening a patient for cancer or precancer, the method comprising the steps of:

detecting in a patient tissue or body fluid sample comprising exfoliated cells and cellular debris, nucleic acid fragments that are greater than 200 base pairs in length;

the presence of said fragments being a positive screen for cancer or precancer.

2. The method of claim 1, wherein said detecting step comprises conducting an amplification reaction designed to amplify only nucleic acids in said sample that are greater than 200 base pairs in length.

3. The method of claim 1, wherein said sample is selected from the group consisting of stool, pus, and urine.

4. The method of claim 1, further comprising the step of enriching said sample for human DNA.

5. The method of claim 1, further comprising the step of isolating human DNA from said sample.

6. A method for screening a patient for cancer or precancer, the method comprising the steps of:

determining in a patient tissue or body fluid sample a first amount of nucleic acid fragments greater than 200 base pairs in length;

determining in said sample a second amount of nucleic acid fragments less than about 200 base pairs in length;

determining a ratio between said first amount and said second amount; and

identifying a positive screen if said ratios exceeds a threshold ratio for patients who do not have cancer or precancer.

7. A method for screening a patient for cancer or precancer, the method comprising the step of

detecting in a patient tissue or body fluid sample comprising exfoliated cells a nucleic acid fragment of a length that is not expected to be present in said sample in a healthy patient;

the presence of said fragment being a positive screen.

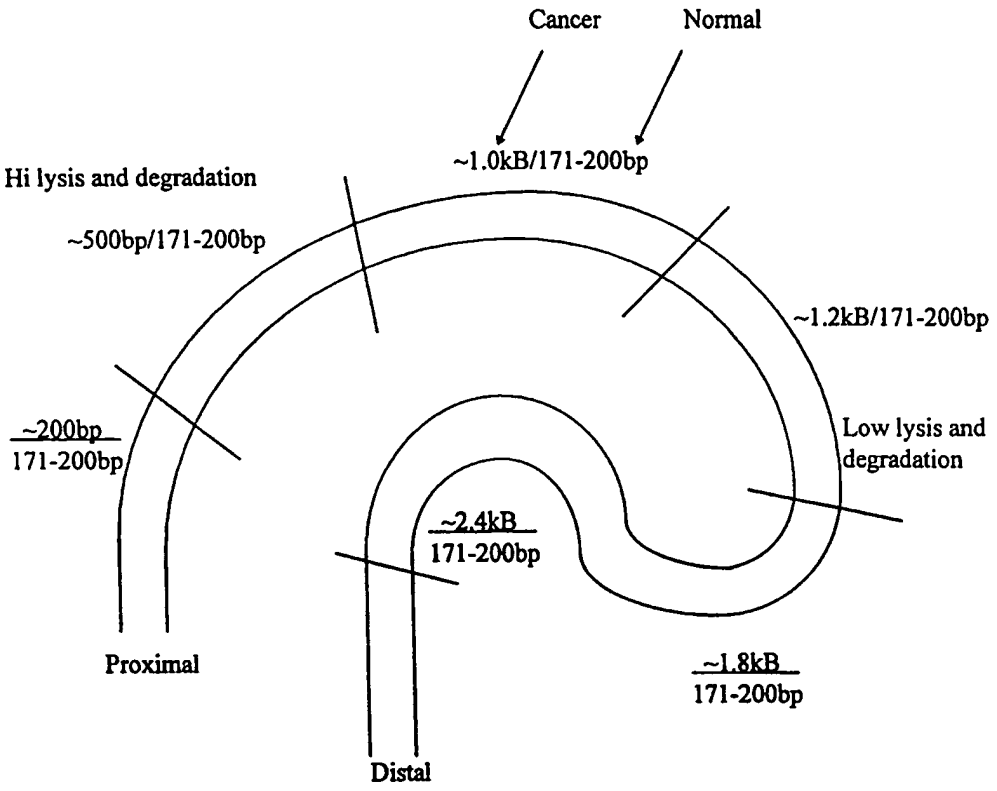
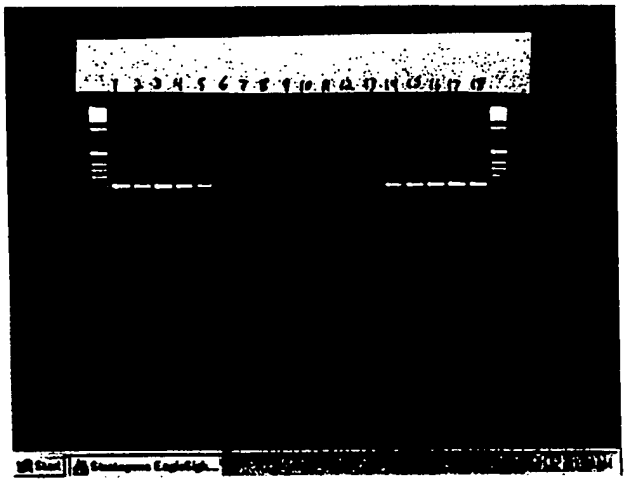

Cancer
Normal
~1.0kB/171-200bp
Hi lysis and degradation
~500bp/171-200bp
~1.2kB/171-200bp
~200bp
171-200bp
Low lysis and degradation
~2.4kB
171-200bp
Proximal
~1.8kB
171-200bp
Distal

Figure 1

Figure 2

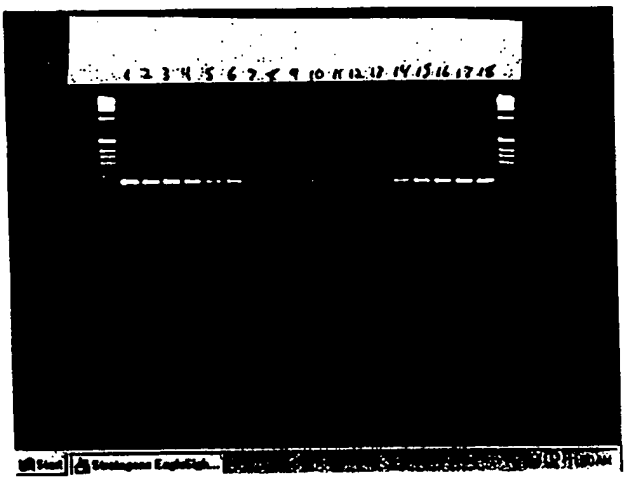


200bp amplifications
33 Cycles

Lane	Q#	Sample Type	Sample Number	Grade
1	7903.8	Abnormal	1	A
2	5627.4	Abnormal	2	A
3	8809.11	Abnormal	3	A
4	5421.94	Abnormal	4	A
5	1838.07	Positive Control		B
6	-549.23	Normal	5	C
7	-715	Normal	6	C
8	-1605.13	Normal	7	C
9	-824.73	Normal	8	C
10	259.77	Normal	9	C
11		Neg Control	-	
12		Neg Control	-	
13	400	400	Standard	
14	2000	2000	Standard	
15	4000	4000	Standard	
16	6000	6000	Standard	
17	8000	8000	Standard	
18	10000	10000	Standard	

A= >2000
B= 500-2000
C= <500

Figure 3

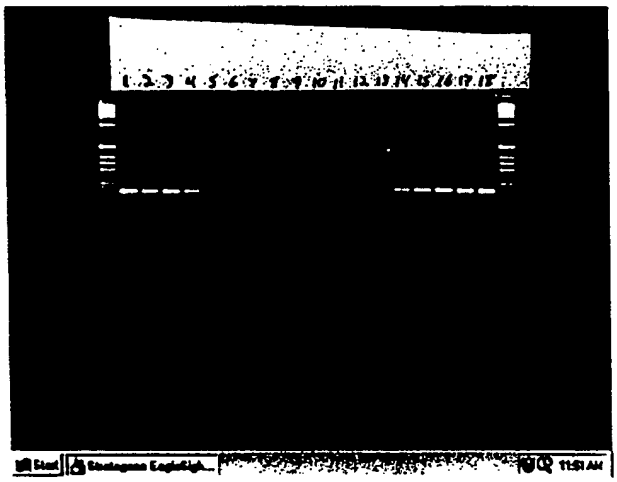


200bp amplifications
35 Cycles

Lane	Q#	Sample Type	Sample Number	Grade
1	10851.04	Abnormal	1	A
2	8862.34	Abnormal	2	A
3	9777.85	Abnormal	3	A
4	6874.28	Abnormal	4	A
5	2392.07	Positive Control		B
6	3080.62	Normal	5	B
7	813.45	Normal	6	C
8	-720.04	Normal	7	C
9	-442.2	Normal	8	C
10	1353.86	Normal	9	B
11		Neg Control	-	
12		Neg Control	-	
13	400	400	Standard	
14	2000	2000	Standard	
15	4000	4000	Standard	
16	6000	6000	Standard	
17	8000	8000	Standard	
18	10000	10000	Standard	

A= >5000
B= 1000-5000
C= <1000

Figure 4

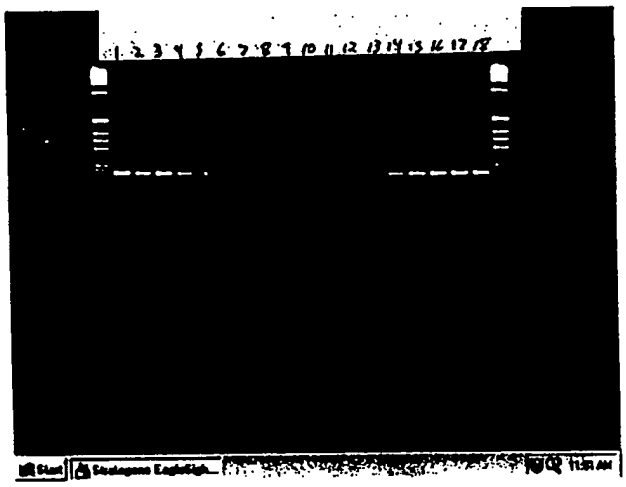


200bp amplifications
34 Cycles

Lane	Q#	Sample Type	Sample Number	Grade
1	8428.34	Abnormal	1	A
2	4917.31	Abnormal	2	A
3	7742.22	Abnormal	3	A
4	3049.85	Abnormal	4	A
5	409.5	Positive Control		B
6	-682.75	Normal	5	C
7	-781.09	Normal	6	C
8	-1099.28	Normal	7	C
9	-1015.39	Normal	8	C
10	359.74	Normal	9	B
11		Neg Control	-	
12		Neg Control	-	
13	400	400	Standard	
14	2000	2000	Standard	
15	4000	4000	Standard	
16	6000	6000	Standard	
17	8000	8000	Standard	
18	10000	10000	Standard	

A= >750
B= 250-750
C= <250

Figure 5

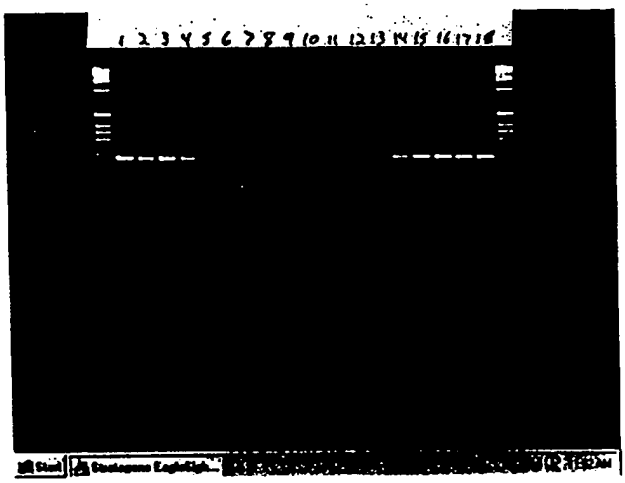


200bp amplifications
33 Cycles

Lane	Q#	Sample Type	Sample Number	Grade
1	7879.15	Abnormal	1	A
2	4079.09	Abnormal	2	A
3	7995.95	Abnormal	3	A
4	2600.3	Abnormal	4	A
5	1698.19	Positive Control		B
6	-405.32	Normal	5	C
7	-466.15	Normal	6	C
8	-1046.47	Normal	7	C
9	-764.83	Normal	8	C
10	105.05	Normal	9	C
11		Neg Control	-	
12		Neg Control	-	
13	400	400	Standard	
14	2000	2000	Standard	
15	4000	4000	Standard	
16	6000	6000	Standard	
17	8000	8000	Standard	
18	10000	10000	Standard	

A= >2000
B= 500-2000
C= <500

Figure 6

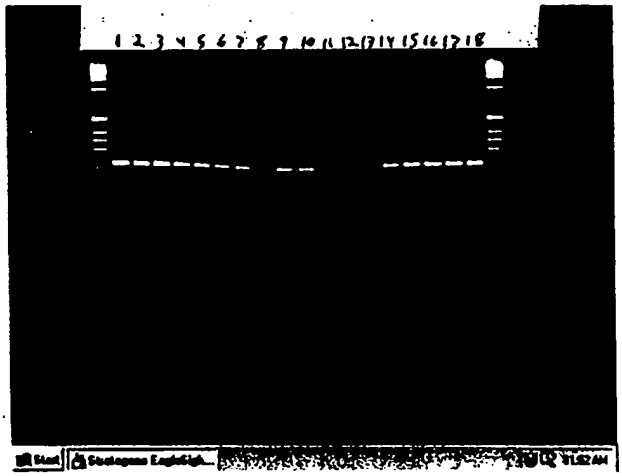


200bp amplifications
34 Cycles

Lane	Q#	Sample Type	Sample Number	Grade
1	7852.95	Abnormal	1	A
2	4797.07	Abnormal	2	A
3	8543.47	Abnormal	3	A
4	3597.23	Abnormal	4	A
5	943.84	Positive Control		B
6	-296.7	Normal	5	C
7	-5.48	Normal	6	C
8	-896.94	Normal	7	C
9	-196.87	Normal	8	C
10	414.81	Normal	9	C
11		Neg Control	-	
12		Neg Control	-	
13	400	400	Standard	
14	2000	2000	Standard	
15	4000	4000	Standard	
16	6000	6000	Standard	
17	8000	8000	Standard	
18	10000	10000	Standard	

A= >2000
B= 500-2000
C= <500

Figure 7

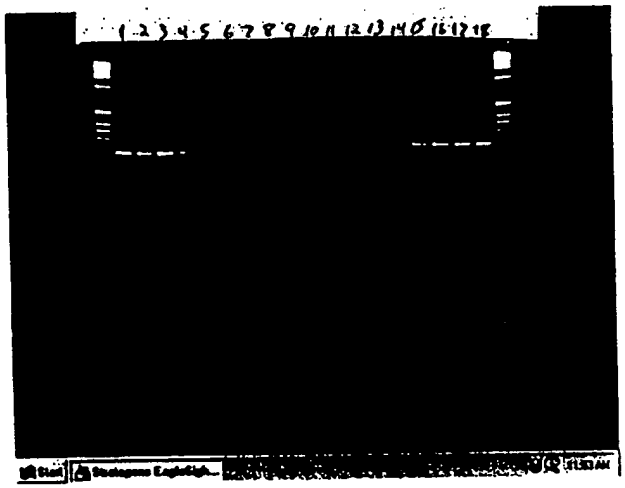


200bp amplifications
34 Cycles

Lane	Q#	Sample Type	Sample Number	Grade
1	7660.6	Abnormal	1	A
2	7032.89	Abnormal	2	A
3	8364.31	Abnormal	3	A
4	6892.04	Abnormal	4	A
5	4883.47	Positive Control		A
6	1934.67	Normal	5	B
7	1380.84	Normal	6	B
8	-964.17	Normal	7	C
9	1729.51	Normal	8	B
10	2221.69	Normal	9	B
11		Neg Control	-	
12		Neg Control	-	
13	400	400	Standard	
14	2000	2000	Standard	
15	4000	4000	Standard	
16	6000	6000	Standard	
17	8000	8000	Standard	
18	10000	10000	Standard	

A= >5000
B= 1000-5000
C= <1000

Figure 8



200bp amplifications
33 Cycles

Lane	Q#	Sample Type	Sample Number	Grade
1	8519.13	Abnormal	1	A
2	5745.19	Abnormal	2	A
3	9765.65	Abnormal	3	A
4	4153.79	Abnormal	4	A
5	1869.33	Positive Control		B
6	418.37	Normal	5	C
7	405.91	Normal	6	C
8	-258.08	Normal	7	C
9	141.64	Normal	8	C
10	450.78	Normal	9	C
11		Neg Control	-	
12		Neg Control	-	
13	400	400	Standard	
14	2000	2000	Standard	
15	4000	4000	Standard	
16	6000	6000	Standard	
17	8000	8000	Standard	
18	10000	10000	Standard	

A= >2000
B= 500-2000
C= <500

Figure 9

1.8 kb amplifications
36 Cycles

Lane	Q#	Sample
1		Neg Control
2	102.935	Abnormal
3	260.645	Abnormal
4	0.075	Normal
5	48.305	Abnormal
6	0.045	Normal
7	18.575	Normal
8		Neg Control
9		Neg Control
10	75	75
11	125	125
12	250	250
13	500	500
14	1000	1000

Abnormal / Normal cutoff 40

Figure 10

1.8 kb amplifications
38 Cycles

Lane	Q#	Sample
1		Neg Control
2	81.84	Abnormal
3	91.515	Abnormal
4	0.04	Normal
5	24.86	Abnormal
6	0.88	Normal
7	17.25	Normal
8		Neg Control
9		Neg Control
10	75	75
11	125	125
12	250	250
13	500	500
14	1000	1000

Abnormal / Normal cutoff 20

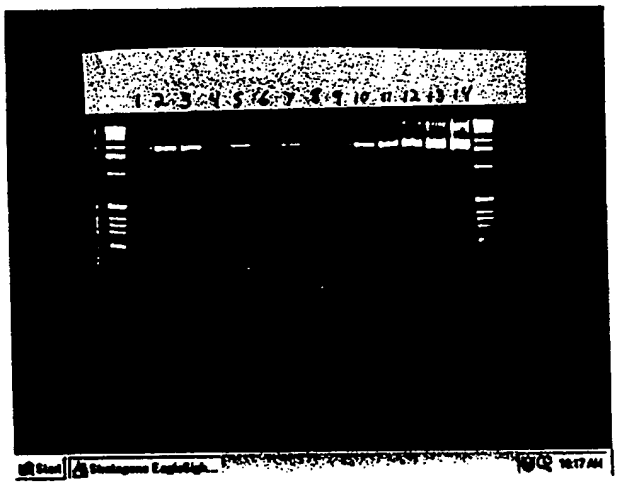


Figure 11

1.8 kb amplifications
40 Cycles

Lane	Q#	Sample
1		Neg Control
2	70.72	Abnormal
3	92.78	Abnormal
4	96.76	Abnormal
5	0.00	Normal
6	29.85	Abnormal
7	0.00	Normal
8	2.00	Normal
9		Neg Control
10		Neg Control
11	75	75
12	125	125
13	250	250
14	500	500
15	1000	1000
16	2000	2000

Abnormal / Normal cutoff 10